

# Análise do Movimento Humano por Videogrametria

**Anita Maria da Rocha Fernandes**  
anita.fernandes@univali.br  
UNIVALI

**Eros Comunello**  
eros.com@univali.br  
UNIVALI

**Fábio Duarte Machado**  
fabio.d.m@univali.br  
UNIVALI

**Resumo:** A análise do movimento do corpo humano é objeto de estudo de diferentes áreas, tais como medicina e esporte. Grande parte deste tipo de análise é realizada com o auxílio de recursos fotográficos ou videográficos, ou seja, de uma forma ou de outra a imagem está acessível sempre que necessária, desde que seja devidamente armazenada. Com o grande avanço tecnológico, as imagens deixaram de ser somente armazenadas em negativos ou papel especial, e tal como as imagens de vídeo, passaram a ser armazenados em formato de arquivo digital. Uma vez neste formato, as imagens podem ser processadas por meios computacionais com a aplicação de técnicas de processamento digital de imagens. Este trabalho visou a pesquisa e concepção de um protótipo de software para análise do movimento humano, baseando em sequências de imagens digitais de forma estereoscópica, no qual se utiliza técnicas de segmentação de imagens, morfologia matemática, rastreamento de objetos e predição de localização. O processamento ocorre em pelo menos dois vídeos, quantidade necessária para a realização da reconstrução tridimensional dos objetos rastreados, também conhecidos no campo da biomecânica por marcas técnicas. Objetivou-se realizar a localização e o rastreamento das marcas técnicas de forma automatizada e o tratamento de sobreposições e oclusões aplicando recursos computacionais preditivos. Todas as coordenadas obtidas das respectivas marcas técnicas são armazenadas em um primeiro momento em memória e posteriormente armazenadas em forma de arquivo de coordenadas, para que possa ser utilizado por outro software de reconstrução tridimensional.

**Palavras Chave:** Processamento digita - Rastreamento de obje - Videogrametria - -

## 1. INTRODUÇÃO

O movimento do corpo humano apresenta alta complexidade, dado ao fato que este apresenta variações, tanto na comparação intra-sujeito, quanto na comparação inter-sujeitos, variações estas possibilitadas por suas articulações e seus graus de liberdade (ARAUJO *et al.*, 1980). Segundo Paul (1995), o pioneiro na realização de uma ampla análise do movimento de partes do corpo humano em três dimensões foi Fischer em 1898, que fez uso de quatro câmeras em um ambiente escuro, e registrou o deslocamento de um sujeito vestindo roupa escura, tendo fixado a ela tubos luminosos, representando assim, partes do corpo em sucessivas posições. Embora todo o procedimento experimental seja extremamente dispendioso, a análise dos resultados pode ser caracterizada como completa e integrada. Completa porque permitiu a aquisição de dados de todos os segmentos corporais, e integrada pela utilização de parâmetros antropométricos e cinemáticos concomitantemente.

O princípio da análise do movimento do corpo humano através das imagens consiste na captura de uma sequência de imagens, por meio de fotos (fotogrametria), ou vídeo (videogrametria), com o auxílio de uma ou mais câmeras, obtendo dados do objeto de interesse através das medidas realizadas nessas imagens (LAGO, 1999). Para formar segmentos e referências para a realização das medidas do objeto de interesse, as imagens dos vídeos utilizados como base para análises de movimento fazem uso de pontos, ou marcadores reluzentes, permitindo assim uma melhor diferenciação dele em relação ao ambiente. Entretanto, em alguns sistemas que fazem uso de marcadores e imagens de vídeo, o correlacionamento dos marcadores é realizado manualmente, quadro a quadro, o que torna o processo de análise massivo. Para se ter uma noção deste processo, toma-se como base, um vídeo com 5 segundos de duração, contendo 1.000 *frames*/segundo e 16 marcadores por vídeo, como mostra a Figura 1, e realiza-se a multiplicação entre eles, obtendo assim, um total de 240.000 cliques de mouse.



**Figura 1:** Filmagem simultânea das marcas técnicas.

Sendo assim, a aplicabilidade de conceitos, métodos e técnicas de processamento de imagens digitais para a realização da automação da detecção e rastreamento dos marcadores (*tracking*), melhora a precisão e a velocidade de obtenção dos resultados para a análise. Dentro deste contexto, este artigo apresenta o desenvolvimento de sistema computacional capaz de analisar o movimento humano por videogrametria.

## 2. REVISÃO DA LITERATURA

Esta seção visa apresentar os principais fundamentos, técnicas de processamento de imagens digitais, algoritmos que compõem o rastreamento de objetos e modelagem em três dimensões, utilizados na base de desenvolvimento do sistema.

### 2.1. IMAGEM

Em se tratando de imagem, ela pode ser definida como uma função bidimensional  $f(x,y)$ , onde  $x$  e  $y$  são coordenadas espaciais (plano), e a amplitude de  $f$  em qualquer para de coordenadas  $(x,y)$  é chamada de intensidade ou nível de cinza da imagem naquele ponto (GONZALEZ & WOODS, 2008). Definida no “mundo real”, imagem é considerada uma função de duas variáveis reais,  $f(x,y)$ , na qual a amplitude (brilho) é dada na coordenada real de posição  $(x,y)$ , e pode ser assim considerada por conter sub-imagens, sendo as vezes referida como regiões de interesse (ROI's, do inglês *Region of Interest*), ou simplesmente regiões. Este conceito reflete o fato de que imagens frequentemente contêm coleções de objetos, e cada qual pode ser base para uma região (GERBRANDS, VLIET & YOUNG, 1998).

Uma imagem consiste de um número infinito de pontos, com valores contínuos de coordenadas e amplitudes. Para converter esta imagem para um formato digital, coordenadas e amplitudes devem ser discretizadas, ou seja, os pontos da imagem serão convertidos para *pixels*, enquanto as amplitudes em valores discretos. De acordo com Gonzalez e Woods (2008), imagem digital é composta por um número finito de elementos, cada qual com um determinado local e valor discreto, identificados pelos índices das linhas e colunas da matriz que a representa.

. Para analisar uma imagem em sua completude, faz-se necessário compreender a cor e sua percepção. Segundo Stones (2003), a visão das cores começa com a luz que atinge a retina, órgão fotossensível e que está presente na parte de trás do olho, e a partir da codificação da luz em sinais, realizada por células especiais chamadas cones, transmite-os ao cérebro que por sua vez interpreta estes sinais e fornece a percepção de cor.

Quando a luz entra nos olhos como um espectro de cores, distribuídos por comprimento de onda, afeta a retina, na qual é composta por dois tipos de receptores, cones e bastonetes. Os seres humanos possuem três tipos de cones, longo, médio e curto (*Long, Medium, Short* ou LMS), de acordo com o comprimento de onda que absorvem, ou mais informalmente, *red* (vermelho), *green* (verde) e *blue* (azul) (RGB) (STONES, 2003). Os bastonetes são receptores longos e finos, e os cones apresentam geralmente uma estrutura mais curta e mais grossa. Ambos possuem importantes distinções operacionais, os bastonetes são mais sensíveis que os cones à luz. Em baixos níveis de iluminação, os bastonetes provêm uma resposta visual chamada visão escotópica (*scotopic vision*). Cones respondem a maiores níveis de iluminação, sua resposta é chamada de visão fotópica (*photopic vision*) (PRATT, 2007).

As cores percebidas são divididas em duas classes: cromática e acromática. Cromática são as cores que variam do branco ao preto. As acromáticas, na qual se percebe nas superfícies dos objetos, pode ser caracterizada por três componentes (ABIDI & KOSCHAN, 2008): Matiz: descreve o tipo de cromaticidade que uma cor possui, e é indicada geralmente com palavras como vermelho, amarelo e azul, e representa a cor dominante; Saturação: descreve a pureza de uma cor, ou a medida do grau em que a cor pura é diluída pela luz branca. Quanto mais claro for, mais saturada é a cor; e Luminosidade ou brilho: indica a força da sensibilidade à luz, pois está ligada a sensibilidade de cada cor. Quanto maior a intensidade da iluminação, mais clara a cor aparece. Cromaticidade descreve a dissimilaridade de uma cor para uma cor acromática de luminosidade igual, ou seja, uma luz igual (ABIDI & KOSCHAN, 2008).

## 2.2. ESPAÇOS DE CORES

Espaço de cores é um espaço geométrico  $m$ -dimensional com eixos adequadamente definidos de modo que os símbolos de todas as percepções de cores possíveis aos seres humanos ou a outros animais se encaixam em uma ordem correspondente à ordem psicológica. Neste espaço cada percepção da cor é representada como um ponto. As representações simbólicas da percepção da cor neste espaço formam a cor sólida. As primeiras propostas para cores sólidas tinham formas geométricas simples: dupla pirâmide triangular, esfera, cone, e assim por diante (KUEHNI, 2003).

**Modelo RGB:** este modelo se baseia na mistura e nos comprimentos de onda internacionalmente normalizados das três cores primárias, vermelho, verde e azul. Neste espaço uma imagem colorida é tratada como uma função de vetor com três componentes, e estes determinados pela medida de intensidade da área de luz visível nos três comprimentos de onda. Para uma imagem digital colorida  $C$ , apresentando três canais, os três componentes do vetor aparecem para cada pixel da imagem  $(x, y)$  (ABIDI & KOSCHAN, 2008).

**Modelo HSB:** também conhecido por HSV (Hue, Saturation, Value), o modelo HSB (Hue, Saturation, Brightness) é particularmente comum no domínio da computação gráfica. Assim como no espaço de cor HSI (Hue, Saturation, Intensity), matiz, saturação e brilho são usados como eixos de coordenadas (ABIDI & KOSCHAN, 2008).

**Modelo CMYK:** diferentemente de um sistema aditivo de cores RGB, este modelo faz uso de um esquema de cores subtrativas, no qual cada cor impressa reduz a intensidade da luz no local da impressão. A impressão de cores requer no mínimo três cores primárias; tradicionalmente ciano (C), magenta (M) e amarelo (Y) devem ser usadas (BURGER & BURGER, 2009). Na prática, segundo Burger e Burger (2009), a cor produzida pela saturação total das três cores não é fisicamente um preto verdadeiro. Entretanto, as três cores primárias, ciano, magenta e amarelo, são normalmente suplementadas com a cor preta (K) para aumentar a gama de cores.

## 2.3. VIDEO

O vídeo utiliza-se de umas das propriedades do olho humano chamada de persistência retiniana ou retenção retiniana, a qual provoca a ilusão de movimento. A persistência de uma imagem detectada pelo olho humano tem duração de 1/12 de segundos. De acordo com Pratt (2007), a taxa de quadros, nome que designa a sequência de imagens reproduzidas, para causar a sensação de continuidade no tempo deve ser maior que 15 quadros por segundo. Para Jack (2001), embora um vídeo pareça um movimento contínuo, na verdade é uma série de imagens estáticas mudando rápido o suficiente parecendo um movimento contínuo. O movimento de cada *pixel* em um vídeo é muitas vezes chamado fluxo óptico (*optic flow*) e é medido pela tentativa de encontrar os *pixels* na imagem seguinte, que correspondem a um *pixel* da imagem atual (correspondência a ser medida pela similaridade em cor, intensidade e textura). Em princípio, há um vetor de fluxo óptico para cada *pixel*, formando um campo de movimento (FORSYTH & PONCE, 2002).

Segundo Tekalp (1995), quase todos os sistemas de vídeo digital fazem uso de componentes de representação para o sinal de cor, assim como a maioria das câmeras de vídeo à cores que fornecem saídas RGB, que são digitalizadas individualmente. Em vídeo digital, não há a necessidade do uso de pulsos de sincronismo, como no vídeo analógico, desde que o computador conheça exatamente onde começa uma nova linha, e a sua quantidade de *pixels*.

Normalmente sistemas de tempo real, são baseados em normas de vídeo, o que significa que a aquisição das imagens, bem como a conversão dos dados digitais em um sinal de vídeo tem que se conformar com as normas internacionais, tais como (ERHARDT-FERRON, 2000): CCIR: Comité Consultatif International des Radiocommunications, Europa; RS-170: definida pela Electronics Industries Association (EIA), USA; PAL: Phase



Alternation Line, baseada na CCIR; SECAM: Sequencial Couleur á Memoire, baseada na CCIR.

#### 2.4. RESOLUÇÃO

De acordo com Poynton (2003), em computação o termo resolução, infelizmente passou a referir-se apenas à contagem de *pixels* existentes na matriz de *pixels*, sem levar em conta a sobreposição na captura, ou sobreposição na exibição, o que pode levar ao decremento da quantidade de detalhes da imagem. Seguindo esta mesma lógica, resolução diz respeito ao número máximo de pares de linhas ou ciclos, que podem ser resolvidos pela tela do monitor, não havendo relacionamento com a nitidez percebida.

Para Jack (2005), as resoluções que se destacam são: Standard Definition: usada pelos sistemas NTSC, PAL e SECAN e apresenta uma matriz de 720x480 ou 720x576 entrelaçadas; Enhanced Definition: diferencia-se da Standard Definition por apresentar uma resolução de 720x480 ou superior de forma progressiva e não entrelaçada; High Definition: apresenta resolução 1920x1080 entrelaçada ou 1280x720 progressiva; e Full High Definition: apresenta resolução 1920x1080 progressiva.

#### 2.5. SINCRONIZAÇÃO DE VÍDEOS

Atualmente muitas aplicações fazem uso de captura de eventos por múltiplas câmeras, tais como sistemas de segurança e reconstrução 3D para cenas dinâmicas de roteiros de cinema. Um dos problemas enfrentados nesses tipos de aplicações é a sincronização que se refere a S sequências de diferentes vídeos que se sobrepõem no tempo; identificar um quadro de cada uma das diferentes sequências que se referem ao mesmo ponto no tempo. O problema da sincronização pode ser expresso em encontrar um conjunto de quadros de números  $f_1, f_2, \dots, f_s$ , um de cada sequência, de tal forma que a igualdade de sincronização seria  $T_1(f_1) = T_2(f_2) = \dots = T_s(f_s)$  (BOSE, LAGANIERE & WHITEHEAD, 2010).

O uso de soluções baseadas em hardware é uma das maneiras para lidar com a sincronização, onde um equipamento especial insere um código único a cada quadro de vídeo, e os quadros que apresentam o mesmo código correspondem ao mesmo instante de tempo. Este tipo de solução é frequentemente complicado, e não é viável quando as câmeras estão fisicamente distantes. Outra solução é usar a visão computacional e aplicar similaridade de medidas de quadros, e histograma de cores. (SUJIT, JAWAHA & NARAYANAN, 2008)

### 3. PROCESSAMENTO DIGITAL DE IMAGENS

Processamento Digital de Imagens (PDI) é a tecnologia de aplicação de uma série de algoritmos de computação para processar imagens digitais. Os resultados deste processo podem ser imagens, um conjunto de características representativas ou as propriedades da imagem original (ZHANG, ZHOU & WU, 2010). Para entender em linhas gerais o PDI, é necessário compreender o que é morfologia matemática e segmentação de imagens.

Segundo Soille (2003), morfologia matemática (MM) ou simplesmente morfologia pode ser definida como uma teoria para a análise de estruturas espaciais. É chamado de morfologia, porque tem como objetivo analisar o conteúdo e a forma dos objetos. Matemática, no sentido de que a análise é baseada na teoria dos conjuntos, geometria integral e álgebra de grade. Morfologia matemática não é apenas uma teoria, mas também uma poderosa técnica de análise de imagem.

Conforme Pratt (2007), o processamento de imagem de forma morfológica é um tratamento em que a forma espacial ou a estrutura do objeto dentro de uma imagem são modificadas, e contêm como operações fundamentais a dilatação, erosão e esqueletização. De acordo com Jähne (2002), quase todas as operações de processamento de imagem visam um melhor reconhecimento dos objetos de interesse, e um dos passos é analisar cada *pixel*

individualmente para verificar se ele pertence ao objeto de interesse ou não. Esta operação é chamada de segmentação, e produz uma imagem binária. Um *pixel* tem valor 1 se pertence ao objeto, caso contrário é 0.

Pratt (2007) explica que a segmentação de uma imagem implica na divisão ou separação da imagem em regiões com atributo similar. O atributo mais básico para a segmentação é a amplitude da luminosidade para imagens monocromática e componentes de cor para imagens coloridas. Bordas e textura também são atributos úteis para a segmentação.

#### **4. ANÁLISE DE MOVIMENTO**

Segundo Boyle, Hlavac e Sonka (2009), o interesse em processar o movimento tem aumentado nos últimos anos, devido aos avanços das metodologias de análise e a capacidade de processamento. Novos aplicativos são motivados pelas necessidades de detecção e rastreamento da face humana, detecção de movimento de pedestres e veículos, ou ainda na assistência a condutores, veículos autônomos, navegação de robôs e interface de usuário, etc.

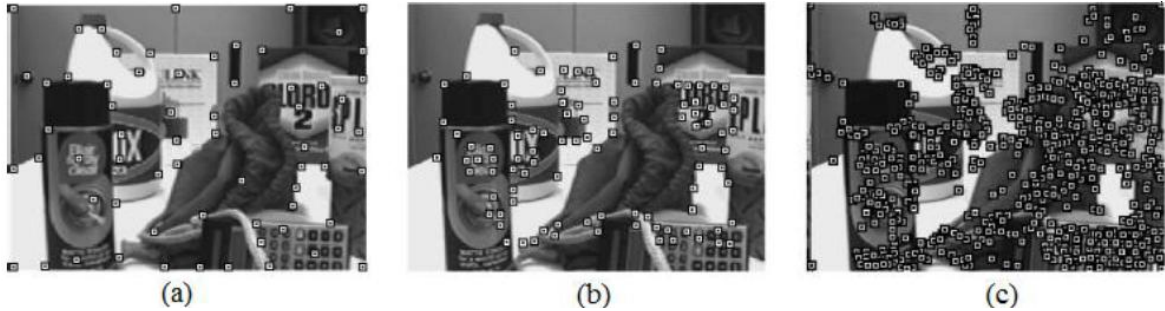
A análise de movimento envolve a interpretação de dados de imagem ao longo do tempo. É crucial para uma série de tarefas de visão, tais como detecção de obstáculos, estimativa de profundidade, análise de vídeo, interpretação de cenário, compressão de vídeo, etc. Esta análise é difícil porque requer a modelagem das relações complexas entre os dados da imagem observada e o movimento dos objetos, e os padrões de movimento na cena visual (XUE & ZHENG, 2009). Os conceitos básicos que envolvem a análise de movimento são:

**Deteção de Objetos**: de acordo com Pratt (2007), uma das maneiras mais fundamentais de detecção de objetos dentro de um campo de imagem é o casamento de modelos (do inglês *template matching*), em que uma réplica do objeto de interesse é relacionada a todos os objetos desconhecidos no campo da imagem. Se a relação entre o modelo e o objeto desconhecido for suficientemente próxima, o objeto desconhecido é rotulado como objeto modelo.

**Oclusão**: de acordo com Grades (2005), oclusão significa que uma parte de um objeto é escondida por outro objeto. Em outras palavras, a região escondida de uma imagem é uma parte espacialmente coerente que não pode ser vista neste objeto. As regiões escondidas geram descontinuidades em profundidade, que são importantes para qualquer processo em que se necessita conhecer os limites do objeto, tais como reconhecimento de padrões, segmentação e análise de movimento, e identificação de objetos. Sendo assim um dos atuais obstáculos em muitos campos científicos como a visão computacional, computação gráfica e investigação cerebral (GRADES, 2005).

##### **4.1. ALGORITMOS DE RASTREAMENTO E PREDIÇÃO**

Segundo Javed, Yilmaz e Shah (2006), os detectores pontuais são utilizados para localizar pontos de interesse em imagens que têm uma textura expressiva em suas respectivas localidades. Pontos de interesse têm sido muito utilizados no contexto do movimento, na visão estérea, e em problemas de rastreamento (*tracking*). A qualidade desejável de um ponto de interesse é a sua invariância a mudanças na iluminação e ponto de vista da câmera. Na literatura, os comumente detectores de pontos de interesse utilizados incluem o Moravec, Harris, KLT e o SIFT (Figura 2).



**Figura 2:** Detecção de pontos de interesse pela aplicação. (a) de Harris, (b) de KLT, (c) de SIFT.  
**Fonte:** Javed, Yilmaz e Shah (2006).

**Kanade-Lucas-Tomasi (KLT):** propõe um critério de seleção de aspectos que é ótimo por construção porque se baseia de acordo com o funcionamento de métodos de rastreamento e métodos de monitoramento de aspectos, que podem detectar oclusões e características que não correspondem aos pontos do mundo (TOMASI & SHI, 1994).

**SIFT:** Invariant Feature Transform é uma abordagem que transforma uma imagem em uma grande coleção de vetores de características locais, cada um dos quais é invariante a tradução da imagem, escala e rotação, e parcialmente invariante a mudanças de iluminação e projeção 3D ou afins (LOWE, 1999).

**Kalman Filter:** filtro de Kalman é essencialmente um conjunto de equações matemáticas que implementam por meio de recursão, um estimador do tipo “preditor-corretor”. O filtro é robusto em vários aspectos: suporta estimativas dos estados passados, presentes e até mesmo o futuro, e pode fazê-la mesmo quando a natureza precisa do sistema modelado é desconhecida (BISHOP & WELCH, 2001).

#### 4.2. GEOMETRIA DE MÚLTIPLAS VISÕES

O mapeamento de uma cena tridimensional sobre um plano da imagem é uma transformação do tipo muitos-para-um. Isto é, um ponto da imagem não determina unicamente a posição de um ponto correspondente do mundo. Entretanto, a informação de profundidade perdida pode ser obtida pelo uso de técnicas de imageamento estereoscópico (resumidamente, estéreo) (GONZALEZ & WOODS, 2008). A geometria epipolar é a geometria intrínseca projetiva entre dois pontos de vista. É independente da estrutura de cena, e depende apenas de parâmetros internos das câmeras e poses relativas. Esta geometria é geralmente motivada por considerar a busca de pontos correspondentes na correspondência estéreo (HARTLEY & ZISSERMAN, 2003). Qualquer ponto P de uma cena observada por duas câmeras e dois raios correspondentes aos centros ópticos  $C_0$ ,  $C_1$  definem um plano epipolar. Este plano intercepta os planos da imagem nas linhas epipolares (ou apenas epipolares) (BOYLE, HLAVAC & SONKA, 2008).

Para descrever o movimento de uma partícula, é necessário o conhecimento de sua posição no espaço em relação a um dado referencial em função do tempo. A posição da partícula no espaço, relativa ao referencial, é definida com a ajuda de três coordenadas independentes, por exemplo, as coordenadas cartesianas (BARROS *et al.*, 1999). Um dos métodos mais utilizados para realização de calibração de câmera é conhecido como DLT (*Direct Linear Transformation*), ou transformação linear direta, que foi originalmente apresentado por Abdel-Aziz e Karara, 1971. O método DLT utiliza um conjunto de pontos de controle, cujas coordenadas são conhecidas. Os pontos de controle são normalmente fixos a uma estrutura rígida, também conhecida como moldura de calibração (BAI & BARDSLEY, 2007).

A calibração de uma câmera e o conhecimento das coordenadas de um ponto de imagem permite a determinação de uma linha no espaço de forma exclusiva. Se duas câmeras calibradas



observarem a mesma cena do ponto X, suas coordenadas 3D podem ser computadas como a interseção de duas linhas. Este é o princípio básico da visão estéreo, que consiste tipicamente em três etapas (BOYLE, HLAVAC & SONKA, 2008): calibração da câmera; estabelecimento dos pontos correspondentes entre os pares de pontos da imagem esquerda e da imagem direita; reconstrução das coordenadas 3D dos pontos na cena. A precisão da reconstrução 3D depende da disponibilidade e precisão dos dados de configuração da câmera (CYGANNEK & SIEBERT, 2009). A obtenção de coordenadas espaciais de pontos a partir do registro estereoscópico de suas projeções em imagens é uma metodologia bastante difundida na biomecânica, denominada reconstrução tridimensional de coordenadas (BARROS, *et al.*, 1999).

O sistema proposto apenas calcula as coordenadas tridimensionais, a projeção dos pontos graficamente caberá a outro software, tal como AutoCad ou OpenGL.

## 5. TRABALHOS CORRELATOS

Durante o período de pesquisa e análise das necessidades para desenvolvimento do sistema proposto, diversos trabalhos relacionados a processamento de imagens digitais e mais especificamente videogrametria foram encontrados. Entre os mais referenciados e mais próximos do sistema proposto destacam-se o Dvideow e o VIDA.

### 5.1. DVIDEOW

O sistema Dvideow (Digital Video for Biomechanics for Windows 32) foi no Laboratório de Instrumentação para Biomecânica (LIB) da Universidade Estadual de Campinas (Unicamp), com cooperação do Instituto de Computação (IC), Instituto de Física “Gleb Wataghin” (IFGW), ambos da Unicamp, e pelo Institui für Biomechanik-Deutsche Sporthchschule Köln-Deutschland. Ele foi projetado para ser usado em diferentes aplicações, entre as já implementadas estão: análise de marcha (ou caminhada), análise de movimentos respiratórios, rastreamento automático de jogadores de futebol e reconstrução de superfícies do corpo humano (SPORT, 2003; REVISTA FAPESP, 2003).

De acordo com Paterniani (2001), o sistema DVIDEOW é um sistema de análise cinemática tridimensional de movimentos humanos baseado no registro videogramétrico de imagens através de câmeras de vídeo analógicas ou digitais, no qual estão implementados os seguintes procedimentos: captura ou aquisição de imagens de vídeo; medição de pontos nas seqüências de imagens; calibração de câmeras e reconstrução tridimensional dos pontos de interesse. A análise dos segmentos (em movimento ou não) representados a partir dos pontos reconstruídos é realizada através de rotinas desenvolvidas no LIB com o uso dos softwares S-PLUS e MATLAB, as demais implementações em C++.

### 5.2. VIDA

O VIDA (Visualizador Interactivo de Dados biomecânicos Aumentados) é um dos resultados práticos do estudo efetuado por José Manuel Lopes Braz Pereira, em sua dissertação para obtenção do grau de mestre em Engenharia Eletrotécnica e de Computadores pela Universidade Técnica De Lisboa (PEREIRA, 2004).

A idealização do VIDA é decorrente do interesse do Laboratório de Biomecânica da Faculdade de Motricidade Humana (LB/FMH) da Universidade Técnica de Lisboa por aplicação para visualização da atividade muscular de atletas, através de uma solução graficamente intuitiva, e apresenta as seguintes etapas afim de contemplar as expectativas (PEREIRA, 2004): detecção automática, em cada quadro de vídeo proveniente de cada uma das duas câmeras usadas, das marcas iridescentes que assinalam as posições dos centros articulares do indivíduo em andamento; cálculo da posição tridimensional no sistema de coordenadas do mundo das marcas técnicas com base nas posições bidimensionais da imagem



de cada marca em coordenadas dos quadros de vídeo; cálculo da posição e orientação da câmera de vídeo; construção de um modelo virtual do indivíduo com base nos resultados de dois; síntese de uma câmera virtual coincidente com a posição e orientação da câmera calculada em três; projeção do modelo criado em 4 no plano de imagem; e sobreposição da projeção obtida em 6 ao quadro de vídeo.

O sistema de seguimento do atleta e o sistema de segmento do sistema de captação de imagem foram desenvolvidos em MatLab, e o sistema gráfico, módulo de construção de objetos gráficos implementado em C++.

## 6. O SISTEMA PROPOSTO

A interação entre usuário e o sistema proposto é por meio de uma interface gráfica. Esta por sua vez apresenta uma tela principal, como ilustrado na Figura 3, que contém os principais botões de ação, mensagens de sistema, apresentação dos vídeos e base para as demais telas de interação, como abertura dos vídeos, leitura e gravação de arquivos.

À seguir serão apresentados os detalhes de implementação do sistema, considerando os aspectos de exibição dos vídeos, identificação e rastreamento de marcas técnicas; seleção de regiões candidatas, verificação e resolução de conflitos e atualização de marcas técnicas.

O projeto core, gera uma biblioteca que é utilizada na aplicação principal, no qual possui classes de tratamento de exceções e de registro de falhas do sistema (*logs*). O registro das falhas, ou ações do sistema, ocorre com a gravação em arquivo XML do registro da pilha de eventos, ou *stack trace*, ocorridos até o acontecimento do erro, uma atribuição da classe GravaLog. Uma vez que a aplicação é iniciada, um novo arquivo XML é criado, apresentado em seu nome, o nome da aplicação, data de criação e status, como por exemplo, “Messkoffer[08062011204715]DEV.xml.”. Cada método criado, no projeto principal ou nos secundários, é circundado com classes de captura de erro ou exceções, sendo a mais genérica a classe *Exception*, nativa da linguagem JAVA. As demais classes de exceção herdam da classe *Exception*, e são utilizadas de acordo com a funcionalidade que cada método implementa, por exemplo um método que implementa regras de negócio é circundada com a *Exception* e encapsulada por uma exceção do tipo *MesskofferBusException*.

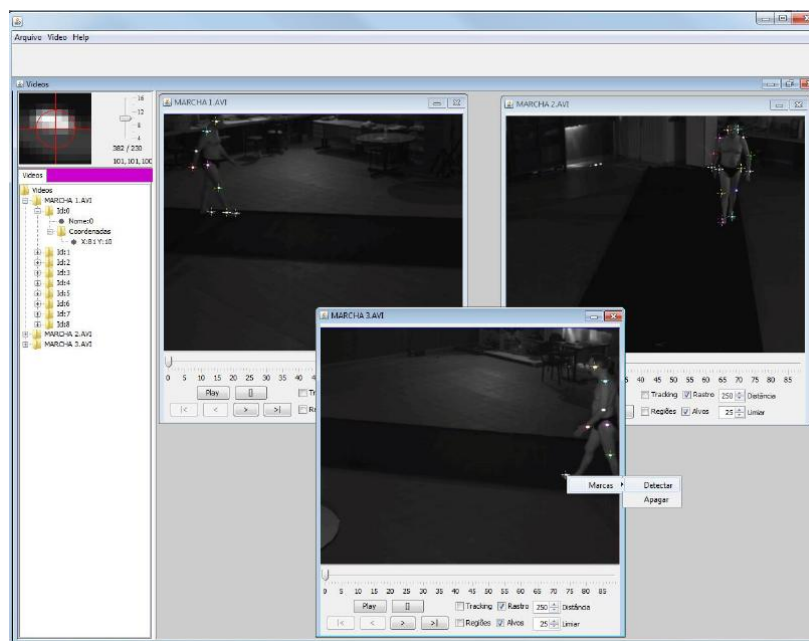


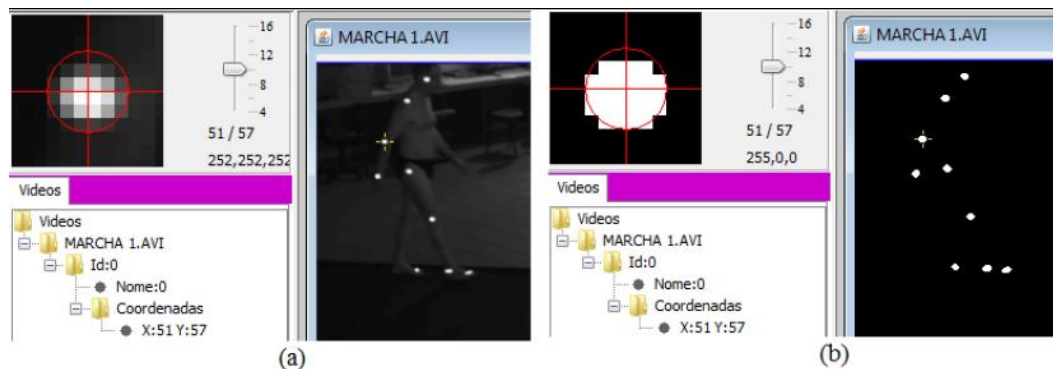
Figura 3: Tela Principal do Sistema.

De modo a obter e visualizar, manipulada ou não, das imagens de vídeo, foi criado uma classe, chamada de “VisualPanel”, que herda métodos da classe *javax.swing.JPanel*, nativa do JAVA e implementa métodos da classe *javax.media.ControllerListener*, nativa da API JMF. A classe “VisualPanel” exerce a função de *player*, com funções básicas como *play* e *stop* além de parâmetros para manipulação das imagens dos frames e rastreamento das marcas técnicas. Selecionado o vídeo, o *player* passa por cinco estágios, definido pela API JMF, de forma a deixar as informações e controladores acessíveis, e ao término, ou seja, estado igual à *Prefetched*, o *player* inicializa variáveis, estado de componentes e a detecção do fundo (*background*).

De modo a obter e visualizar, manipulada ou não, das imagens de vídeo, foi criado uma classe, chamada de “VisualPanel”, que herda métodos da classe *javax.swing.JPanel*, nativa do JAVA e implementa métodos da classe *javax.media.ControllerListener*, nativa da API JMF. A classe “VisualPanel” exerce a função de *player*, com funções básicas como *play* e *stop* além de parâmetros para manipulação das imagens dos frames e rastreamento das marcas técnicas. Selecionado o vídeo, o *player* passa por cinco estágios (Figura 6), definido pela API JMF, de forma a deixar as informações e controladores acessíveis, e ao término, ou seja, estado igual à *Prefetched*, o *player* inicializa variáveis, estado de componentes e a detecção do fundo (*background*). Entre os controladores disponibilizados pela API JMF, dois são indispensáveis para o perfeito funcionamento da captura dos *frames*, *FramePositioningAdapter* e *FrameGrabbingControl*, sendo o primeiro para posicionamento do vídeo no tempo, e o segundo para obter a imagem do *frame*.

## 5.2. IDENTIFICAÇÃO DAS MARCAS TÉCNICAS

A etapa de identificação das marcas técnicas pode ocorrer de duas formas, pela extração da localização da marca técnica através do clique do mouse, auxiliada pela ampliação dos *pixels* arredores do cursor, e com a imagem manipulada, ou de forma automatizada, cabendo ao sistema detectar e definir a localização (ver Figura 4).

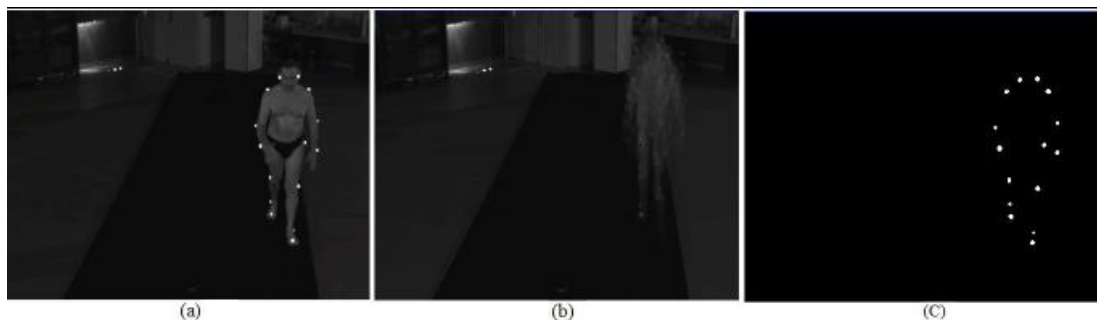


**Figura 4:** Identificação das marcas técnicas. (a) identificação com imagem original. (b) identificação com a imagem manipulada.

Para cada marca técnica localizada um objeto representativo é criado, chamado de “Target”, contendo atributos como, índice, nome, lista das coordenadas percorridas e métodos para auxílio de definição de localização. Os “Targets” são adicionados em uma lista para futura varredura. Assim como o usuário do sistema pode optar em utilizar a imagem manipulada, para melhor identificação da localização das marcas técnicas, o sistema também faz uso deste recurso para detectá-las, entendendo cada região encontrada, como um objeto a ser rastreado. Entende-se como imagem manipulada (Figura 5 (c)) a imagem que passa pelos seguintes processos: subtração da imagem do frame atual pela imagem de fundo; conversão para tons de cinza; aplicação da limiarização; e aplicação da dilatação. O uso da imagem



manipulada se estende durante todo o processo de rastreamento (*tracking*) das marcas técnicas.



**Figura 5:** Imagem manipulada. (a) último frame do vídeo frontal. (b) imagem de fundo ou *background*. (c) subtração da imagem a pela b e dilatada.

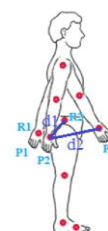
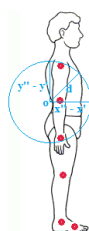
A construção da imagem de fundo é obtida do valor médio de cada *pixel*, processo que ocorre na inicialização do objeto da classe “VisualPanel”, demandando processamento e ocasionando demora na abertura do vídeo.

### 5.3. RASTREAMENTO DAS MARCAS TÉCNICAS

Concluída a etapa de localização das marcas técnicas, o protótipo está apto a iniciar o rastreamento, sempre utilizando a imagem manipulada e os objetos representativos chamados de “Target”. Com o intuito de melhorar a detecção e a seleção das regiões, ou marcas técnicas, dois parâmetros de ajuste foram criados e disponibilizados para alteração, chamados de “Distância” e “Limiar”. Distância corresponde à máxima distância entre a última posição da marca técnica e uma região candidata a avaliação. Se esta estiver numa distância menor que a configurada entra na lista de regiões a serem avaliadas. Limiar: é o valor T utilizado na limiarização, valor que influencia diretamente na formação das regiões. A atribuição de um baixo valor para T pode melhorar a definição de uma região, porém pode causar o aparecimento de regiões indesejáveis.

Outros dois parâmetros, “Rastro” e “Regiões”, também estão disponíveis para alteração, estes referentes à forma de apresentação da imagem. Rastro: habilita ou não o aparecimento do rastro, ou caminho percorrido pelas marcas. Regiões: quando selecionada, apresenta a imagem manipulada. No mesmo painel dos parâmetros de ajuste e visualização encontram-se as funções do *player*, assim como o parâmetro *Tracking*, responsável por habilitar os processos que compõem a etapa de rastreamento.

Para cada região encontrada na imagem manipulada, um objeto representativo é criado, sendo este chamado de *BinaryRegion*, contendo atributos como identificador, quantidade de *pixels* e a coordenada de seu centro. Cada *BinaryRegion* é adicionada a uma lista que é preenchida e limpa a cada nova imagem analisada. Regiões candidatas são as regiões que se encontram dentro da máxima distância configurada, e que possam corresponder a marca técnica no frame subsequente. Tomando como base a Figura 6 e definindo como última posição de uma marca técnica o centro do círculo, têm-se três regiões candidatas a correspondente marca técnica.



**Figura 6:** Seleção de regiões candidatas.**Figura 7:** Exemplo de ocorrência de conflito.

Através da varredura das duas listas, lista de “Target” e lista de “BinaryRegion”, é realizado o relacionamento entre uma “Target” e as regiões candidatas, sendo que este relacionamento é disposto em dois *Map*, tipo de lista que possui relacionamento chave/valor. Um dos *Map* possui como chave a distância entre a última localização da marca técnica e a região candidata, e como valor a própria região candidata, e o segundo *Map* relaciona a “Target”, tendo como chave o seu identificador, com o *Map* anteriormente citado. A escolha deste tipo objeto para guardar os relacionamentos, dá-se ao fato da ordenação pela chave, deste modo, a região que possui a menor distância será a primeira da lista.

Conforme o posicionamento e o deslocamento das marcas técnicas, duas ou mais “Target” podem possuir a mesma região como a primeira de sua lista, inclusive nos casos de oclusão. A ocorrência deste fato ocasiona falha no rastreamento e o registro errôneo da localização da marca técnica. Pode-se observar a representação de um conflito na Figura 7, onde tem-se uma transição de posicionamento acentuada entre a posição P2 e P3, o que faz com que a distância  $d_2$ , entre a última posição P2, da região R1 até R1 na posição P3, seja maior que a distância  $d_1$ , entre R1 na posição P3 e a região R2. Seguindo a lógica de seleção das regiões candidatas, tem-se duas “Target” apontando para a mesma região, a “Target” responsável em rastrear R2 e a “Target” responsável por rastrear R3.

A verificação de conflitos se dá pela comparação uma a uma das “Target”, e as que apresentam como primeira candidata a mesma região são adicionadas a uma nova lista para resolução do conflito. O processo de verificação e resolução de conflito repete-se até que nenhum conflito tenha que ser resolvido, pois a cada resolução um novo pode ser gerado. Após o agrupamento das “Target” em conflito, inicia-se o processo de resolução, ou escolha da melhor região para uma determinada “Target”. Como já explicitado e ilustrado a distância entre a última posição da marca técnica e a região candidata não pode ser decisória para a resolução de um conflito, sendo assim, o conceito de equação da reta, ou área do triângulo é utilizado para auxílio na tomada de decisão. O valor da área do triângulo é obtido pelo valor resultante do cálculo da área do triângulo, com coordenadas  $x$  e  $y$  da penúltima e última localização da marca, e a coordenada central da região candidata. Quanto mais próximo de zero for o valor da área, mais indica que os três pontos avaliados podem compor uma reta. A fim de evitar que se obtenha um valor igual a zero dado o fato que a penúltima e última localização são iguais, é realizada a verificação desta igualdade, e se verdadeira, outra coordenada menos recente e que torne a igualdade falsa deve ser utilizada. Quando se faz necessário o cálculo da área do triângulo, todas as regiões que estão relacionadas a “Target” são avaliadas, não somente a que está em disputa, o que em determinados casos aponta como melhor região uma outra qualquer da lista, diferente da primeira, o que gera mais uma análise na resolução do conflito. A ocorrência deste fato dá-se pelo alinhamento entre as marcas técnicas e independe da distância.

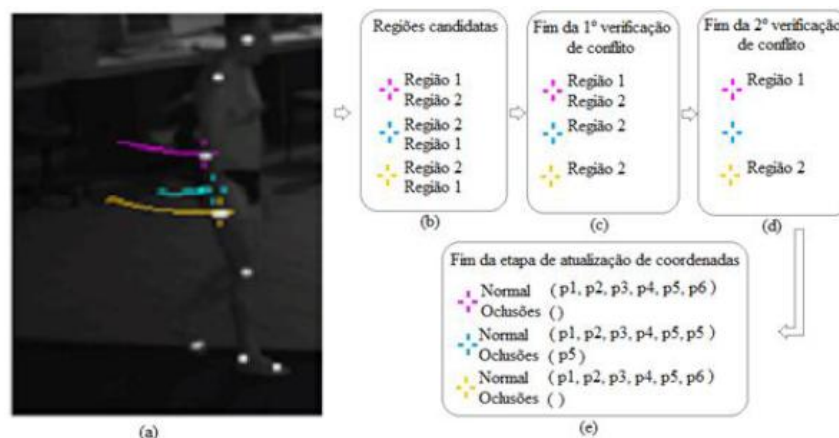
A fim de afinar a melhor escolha da região, dado a ocorrência de conflito, é realizada a comparação entre a multiplicação da área e a distância das duas regiões, a primeira região da lista e a que obteve menor área, e a que apresentar menor valor é definida como melhor região. O uso do cálculo da área do triângulo nem sempre se faz necessário em uma resolução de conflito, pois as “Targets” que apontam para uma mesma região com distâncias bem diferenciadas não são avaliadas pela área do triângulo, o que prevalece é somente a distância. Este valor de diferenciação, ou fator de dúvida, é obtido pela subtração das distâncias. Caso o resultado desta operação seja maior que dez, valor utilizado no sistema conforme análise do vídeo, o cálculo da área do triângulo não é realizado.

Dispostos em um objeto chamado de “ResultDecision”, a área do triângulo, distância e a região candidata são avaliados para a tomada de decisão da melhor região. Dado ao fato que

todas as regiões contidas no map da “Target” são avaliadas, a região candidata atribuída no “ResultDecision” pode ser diferente da região em disputa. Como as duas “Targets” que disputam a mesma região passam pelo cálculo da área do triângulo, geram-se quatro possíveis combinações de retorno: as duas regiões de retorno diferentes da região de disputa e diferentes entre si; as duas regiões de retorno diferente da região de disputa e iguais entre si; as duas regiões de retorno iguais a região de disputa; e as duas regiões de retorno diferente da região de disputa. Ao término de todo o processo de verificação e resolução de conflitos, tem-se nos maps de regiões candidatas nenhuma ou apenas uma região, pois na medida em que se determina a melhor região para uma “Target”, se realiza a exclusão da região de todos maps, exceto do map da “Target” relacionada com a região. A exclusão de uma região das listas, ou maps, de regiões candidatas, também ocorre quando uma “Target” e sua primeira região candidata não possuem conflito, fato comum para as marcas técnicas distante das demais.

A atualização das coordenadas das marcas técnicas respeitam duas condições, a) a de ter uma região relacionada a “Target”, ou uma região em seu *Map*, e b) a condição de *Map* vazio, que caracteriza a existência de oclusão. Na condição em que não há oclusão, simplesmente as coordenadas do centro da região relacionada são adicionadas na lista normal de coordenadas, ou lista de objetos do tipo **java.awt.Point** que possui como atributos coordenada x e y. No caso contrário, existência de oclusão, duas adições são realizadas, uma adição na lista normal e outra na lista de oclusões, sendo que as coordenadas utilizadas são as últimas coordenadas da marca técnica.

Observando-se a Figura 8 tem-se, (a) imagem do frame analisado e rastreamento de três marcas; (b) *Targets* e suas regiões candidatas em ordem crescente de distância entre a última localização da marca e o centro da região; (c) fim da primeira verificação de conflito, onde não ocorre disputa de região; (d) fim da segunda verificação com resolução de conflito, já que duas “Target” disputam a mesma região; (e) lista das localizações das marcas técnicas até o frame analisado, destacando-se as listas da segunda “Target”, em azul, com repetição das coordenadas p5 na lista normal e inclusão na lista de oclusões.



**Figura 8:** Representação dos estados das listas durante o processo de rastreamento.

O uso de uma segunda lista para receber coordenadas oclusas deve-se ao processo de co-relacionamento das marcas técnicas, pois ao se efetuar a reconstrução tridimensional, as coordenadas oclusas podem ser verificadas e substituídas através das coordenadas das outras visões.

## 6. ANÁLISE DOS RESULTADOS

Concluído as principais funcionalidades do sistema proposto, tornou-se possível a realização de uma série de testes permitindo assim avaliar, dentre outros fatores, o alcance dos

objetivos traçados. O carregamento e abertura dos vídeos apesar de ser a etapa que causou preocupações, pois foi considerada como básica e apresentou uma série de problemas, tem desempenho satisfatório, apesar de ser penalizado pelo processo de detecção do fundo da cena. Auxiliada pelo uso da imagem manipulada, a seleção manual das marcas técnicas, efetua satisfatoriamente sua função, de determinar a localização das marcas. A seleção manual das marcas ainda não agrega a função de correção da localização durante o processo de rastreamento, tarefa importante nos casos de sobreposição. O processo de rastreamento, não atingiu a total correção devido ao fato das sobreposições, entretanto um dos principais problemas, a oclusão é contornada de forma satisfatória. Outro fator que vem a satisfazer o rastreamento é o fato que outros objetos de cena não são rastreados como marcas técnicas, problema solucionado com o ajuste dos valores de “Limiar” e “Distância” e subtração do frame atual pelo fundo detectado.

Entre os requisitos funcionais, encontra-se a predição da localização das marcas, contemplado com o uso do *Kalman Filter*, que apresenta bons resultados após conter dados suficientes para ajustar os valores das suas variáveis de controle. Devido à latência de ajuste do filtro, as coordenadas estimadas não são utilizadas, dado o fato que o tratamento de oclusão e a localização de marcas que possuem deslocamentos mais acentuados, tem seu rastreamento prejudicado, algo verificado principalmente nas marcas contidas nos pés. Ao fim do processo de rastreamento de todas as marcas técnicas em todas as visões, utilizando-se os melhores valores de configuração, é possível obter algumas informações importantes.

Considerando-se como erro as marcas que não tiveram suas localizações bem definidas, seja pelo fato da ocorrência de oclusão ou sobreposição, pode-se obter uma porcentagem de erros ocorridos durante todo o processo de rastreamento, sendo este a análise dos três vídeos. Ao realizar-se a soma do número total de localizações, número de frames por número de marcas, chega-se ao valor de 3026 localizações a serem efetuadas, sendo que destas, 136 foram falhas, número de oclusões mais o número de sobreposições, resultando em um percentual de erro de 4,494%. O número de oclusões no vídeo do lado esquerdo se dá as ocorrências de auto-occlusão, braço esquerdo obstruindo a marca técnica localizada na cintura do indivíduo. De acordo com Reis (2009), nas fases iniciais da doença de Parkinson os comprometimentos motores mostram-se mais pronunciados em um dos lados do corpo. Além disso, o indivíduo analisado relatou apresentar os sinais da doença mais pronunciados no seu hemisfério esquerdo. Sendo assim, é esperado que o membro superior esquerdo apresente menor amplitude de movimento e, conseqüentemente, menor velocidade de balanço durante a marcha. Apesar das sobreposições não elevarem, para este caso, o percentual de falha, é imprescindível a busca e implementação de uma solução para esta ocorrência, já que as localizações oclusas podem ser construídas a partir das localizações das outras visões. Desconsiderando o fato que é possível reconstruir a localização de uma marca oclusa, e como dito anteriormente, o tratamento de oclusões foi satisfatória, já que no momento que a marca oclusa volta a aparecer, seu rastreamento é retomado.

## 7. CONCLUSÕES

Durante a realização das etapas de implementação do sistema, uma série de dificuldades surgiram e obtiveram sua devida atenção. Principalmente as que estavam relacionadas as premissas do projeto, e outras sem tanta prioridade, cuja solução foi deixada para trabalhos futuros, nos casos em que o estudo de uma solução demandaria muito tempo. Entre as dificuldades enfrentadas está a incompatibilidade da API JMF com diversos formatos de arquivo de vídeo, instituindo a obrigatoriedade do uso de conversores de vídeo externos, e não ser integralmente compatível com um sistema operacional 64 bits.

A procura de uma alternativa à JMF em trabalhos futuros é fortemente recomendada, com o intuito de deixar o protótipo mais compatível com arquivos multimídia. Outra dificuldade enfrentada e não vencida está a sobreposição das marcas técnicas, caracterizada pelo fato de uma marca técnica transitar para uma coordenada muito próxima a última coordenada de outra marca técnica, e que apresenta o estudo de soluções mais complexas e demanda de maior tempo para sua aplicação. Entende-se que algumas pendências do sistema, como correção manual de localização, co-relacionamento das marcas técnicas e reconstrução tridimensional, podem ser concluídas em trabalhos futuros, uma vez que as premissas do projeto foram realizadas, além de outras funções, como: aplicação de modelos para auxílio na localização das marcas; criação de segmentos entre marcas para determinação de estruturas; e realização de medidas de distâncias entre pontos selecionados com unidades do SI.

## 8. REFERÊNCIAS

- SPORT,** Download: Archiwum. Disponível em: [http://5port.pl/downloads.php?cat\\_id=5&download\\_id=14](http://5port.pl/downloads.php?cat_id=5&download_id=14). Acesso em: 13 de nov. 2010
- ABIDI, M., KOSCHAN, A.** Digital Color Image Processing. New Jersey: John Wiley & Sons. Inc, 2008
- ARAUJO, A. G et al.** Descrição da Movimentação Geral do Corpo, Aplicada á Marcha Patológica. 21º Congresso Brasileiro de Engenharia Biomédica. 2008.
- BAI, L., BARDSLEY D.** 3D Reconstruction Using the Direct Linear Transform with a Gabor Wavelet Based Correspondence Measure. Leeds, UK 2007.
- BARROS, R.M.L. et a.** Desenvolvimento e avaliação de um sistema para análise cinemática tridimensional de movimentos humanos. Revista Brasileira de Engenharia Biomédica, v15, n. 1-2, p. 79-86, jan/ago 1999.
- BISHOP, G., WELCH, G.** An Introduction to the Kalman Filter. University of North Carolina at Chapel Hill, Copyright ACM, Inc, 2001.
- BOSE, P., LAGANIERE, R., WHITEHEAD, A.** Temporal Synchronization of Video: Electronic Letters on Computer Vision and Image Analysis, 2010.
- BOYLE, R., HLAVAC, V., SONKA, M.** Image Processing, Analysis, and Machine Vision: Thomson Learning, 2008.
- BURGER, M., J., BURGER, W.** Principles of Digital Image Processing. Springer-Verlag London Limited, 2009. 71
- ERHARDT-FERRON, A.** Theory and Applications of Digital Image Processing: University of Applied Sciences Offenburg, 2000.
- FORSYTH, D.A., PONCE, J.** Computer Vision: A Modern Approach. New York: Prantice Hall, 2002.
- GERBRANDS, J., VLIET, L.J.; Van YOUNG, I. T.** Fundamentals of Image Processing. Netherlands: Delft University of Technology, 1998.
- GONZALEZ, R. C., WOODS, R. E.** Digital Image Processing, 3/E. New Jersey: Pearson Education Inc., 2008.
- GRADES, E.** Reconstruction of Objects from Images with Partial Occlusion, Dissertation. Ruhr-Universität Bochum, 2005.

- HARTLEY, R., ZISSERMAN, A.** Multiple View Geometry in Computer Vision. New York: Cambridge University Press, 2003.
- JACK, K.** Video Demystified: a handbook for the digital engineer 3rd Edition. LLH Technology Publishing, 2001.
- JÄHNE, B.** Digital Image Processing. Springer-Verlag, 2002.
- JAVED, O., SHAH, M., YILMAZ A.** Object Tracking: A Survey. ACM Computing Surveys, Vol. 38, No. 4, Article 13, 2006
- KUEHNI, R.G.** Color Space and Its Divisions. New Jersey: Wiley-Interscience, 2003.
- LAGO, O. C.** Fundamentos Geométricos e Algébricos da Calibração e Reconstrução Tridimensional. Aplicação na Análise Cinemática de Movimentos Humanos. Dissertação (Mestrado), UNICAMP, Campinas, SP, 1999.
- LOWE, D. G.** Object Recognition from Local Scale-Invariant Features. The Proceedings of the Seventh IEEE International Conference, Vol. 2, 1999, pp. 1150-1157.
- PAUL, J.P.** One hundred years of measurement of human motion. XV Congress of the International Society of Biomechanics. Finland, 1995.
- PATERNIANI, P. E. S.** Desenvolvimento de uma Metodologia para Reconstrução Tridimensional e Análise de Superfícies do Corpo Humano, Dissertação (Mestrado), UNICAMP, Campinas, SP, 2000.
- PEREIRA, J. M. L. B.** Sistema Vida: Uma Visualização Aumentada de Dados Biomecânicos, Universidade Técnica de Lisboa, 2004.
- POYNTON, C.** Digital Video and HDTV - Algorithms and Interfaces. San Francisco: Morgan Kaufmann Publishers, 2003. 73
- PRATT, W.K.** Digital Image Processing. New Jersey: John Wiley & Sons, Inc. 2007.
- REIS, D.** Efeitos da estimulação audiovisual em variáveis motoras relacionadas à marcha de um indivíduo com a doença de Parkinson. Dissertação de Mestrado. Programa de Pós-graduação em Educação Física, UFSC, 2009.
- REVISTA FAPESP,** Detalhes do Movimento. Disponível em: <http://www.revistapesquisa.fapesp.br/?art=2016&bd=1&pg=1&lg=>. Acesso em: 13 de nov. 2010.
- SOILLE, P.,** Morphological Image Analysis –Principles and Applications, 2nd edition: Springer-Verlag, New York, 2003.
- STONES, M.** A Field Guide to Digital Color: A K Peters Ltd., August 2003.
- SUJIT, K., JAWAHAR, C.V., NARAYANAN, P. J.** Video Frame Alignment in Multiple Views. Gachibowli, Hyderabad, 2008.
- SUN.** Java Media Framework API Guide. November,19,1999. Disponível em: <http://www.oracle.com/technetwork/indexes/downloads/index.html>. Acesso em: 9 de jun. de 2011.
- TEKALP, A. M.** Digital Video Processing. University of Rochester, 1995.
- TOMASI, C., SHI, J.** Good Features to Track. IEEE Conference on Computer Vision and Pattern Recognition. Seattle, 1994.
- XUE, J., ZHENG, N.** Statistical Learning and Pattern Analysis for Image and Video Processing. Springer-Verlag, New York, 2009.





**IX SEGeT 2012**

**SIMPÓSIO DE EXCELÊNCIA EM  
GESTÃO E TECNOLOGIA**

Tema: Gestão, Inovação e Tecnologia para a Sustentabilidade

**ZHANG, J., ZHOU, H., WU, J.** Digital Image Processing: Ventus Publishing ApS, 2010.